



PGES
PROFESSIONAL GROWTH AND EFFECTIVENESS SYSTEM

Reliability and Validity

What follows is a brief description of the reliability and validity of the evidences within the Professional Growth and Effectiveness System (PGES), for which reliability and validity data are available.

Definitions

Reliability refers to the degree to which the system, or sources of evidence used to inform judgments within the system, will yield the same result over time.

Validity refers to the degree to which the system, or sources of evidence used to inform judgments within the system, accurately reflect the specific skills or outcomes that are to be measured.

Observation

Kentucky's observation process is based on a principal's deep understanding of how the four domains of the *Kentucky Framework for Teaching* (adapted from the *Charlotte Danielson Framework for Teaching* used in the Measure of Effective Teaching [MET] project) are applied in observation.

The *Framework for Teaching* has been subjected to several validation studies over the course of its development and refinement including an initial validation by the Educational Testing Service (ETS.) Later studies – including one conducted by the Consortium for Policy Research in Education (CPRE) and others assessing the application of the framework for use in teacher evaluation in Cincinnati and Chicago (2011) – have identified small but consistently positive correlations between the framework ratings and student learning outcomes. Validation studies may be found at <http://danielsongroup.org/research/>.

These studies also identified the need for strong observer training systems to ensure consistency of observations. Thus, Kentucky evaluators are required to complete the Teachscape Proficiency Observation Training.

From Kentucky data we can say that raters using the Danielson framework have exhibited inter-rater reliability because they have all been normed using the Teachscape modules. All raters were subjected to the same tests and were required to score the same observations and came up with similar scores.

Student Voice

Kentucky's Student Voice Survey is adapted from the Cambridge Education's Tripod Project surveys that were administered to students in MET project participating classrooms during both years of the study.

Analysis by the MET project finds that teachers' student survey results are predictive of student achievement gains. Further, the MET project finds student surveys produce more consistent results than classroom observations or achievement gain measures.

From Kentucky research, we can say the student voice survey showed high internal consistency and the constructs were all correlated to one another.

Student Growth

Kentucky's Professional Growth and Effectiveness System was designed using the principles that emerged from the findings of the [MET study](#). Culminating findings from the study clearly indicate that:

1. Effective teaching can be measured.
2. Observation ratings can reliably reflect what teachers do throughout the year, as opposed to the subjective impressions of a particular observer or some unusual aspect of a particular lesson.
3. A balanced approach using multiple sources of evidence produces more consistent ratings over time.

The MET study used a quantitative approach to overall teacher ratings. Kentucky's Professional Growth and Effectiveness System, while leveraging the principles of the MET study, relies on a less quantitative approach. Because of the difference in the two models, proving the reliability and validity of Kentucky's overall system will require additional Kentucky-specific data collection. However, given the strength of the research base associated with the instrumentation used to collect the various sources of evidence informing the system, Kentucky can make a strong argument for the face validity of the system. Furthermore, the rationale behind the use of professional judgment in the system provides an even greater argument for the face validity of the system.

Val-Ed 360°

The VAL-ED is a 360° assessment. It is intended to be taken by the principal, the principal's supervisor and all teachers in the school. The VAL-Ed is designed, developed and tested to be both reliable (i.e., provide accurate measurement) and valid (i.e., measure leadership behaviors that lead to improved student achievement). To accomplish these goals, Kentucky followed a multi-stage development process that involved cognitive labs, pilot tests, and field tests. At each stage of the design and development process, the properties of the instrument were investigated through empirical study and expert review. The Kentucky Department of Education (KDE) has completed studies that include: Pilot/Cognitive Interviews; School Pilot Test; Cognitive Interviews of Online Prototype; Bias Review; a spring 2008 Field Test, and Proficiency Standards based on pilot study/field test data.

Ongoing Research Agenda:

By December 2014, KDE should have more data to support the validity of the PGES. Using data from the PGES pilot, KDE intends to see whether there is a correlation between each of the multiple sources of evidence and a teacher's summative rating. Additionally, KDE intends to measure the correlation between each of the sources of evidence and the local and state measures of student growth. To measure the reliability of the classroom observation protocol, KDE will

examine scores using the Danielson instrument and see if there is a high level of internal consistency for that instrument.

In order to measure the reliability of the student voice survey, KDE will monitor the results of the survey to ensure that there is a high level of internal consistency between the items.

Once the PGES is implemented fully, KDE will continue to monitor the reliability and validity of the system. Using the data from the full implementation of PGES, KDE will make changes to the PGES instruments as deemed appropriate. KDE will examine the distribution of the scores of the multiple sources of evidence and the overall performance ratings of teachers. The agency will develop a protocol for adding/modifying questions to the student voice survey and developing questions for students of teachers that have atypical settings (low incidence special education, speech pathologists, etc.). Once multiple years of data are available, KDE will monitor the reliability of the overall performance ratings from year to year. The agency also will continue to monitor the correlation of the sources of evidence with both local and state student growth data. Additionally, KDE will continuously monitor the implementation of the PGES in order to ensure that it is being implemented with fidelity across the state.